UNCLASSIFIED

DRES

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

SR 638

UNCLASSIFIED

High Resolution Liquid Chromatographic Mass Spectrometric Identification of Peptides Using Electrospray Ionization

BY:

J.R. Hancock, P. A. D'Agostino and Lionel Provost
Defence Research Establishment Suffield

and

P. D. Semchuk and R. S. Hodges
Department of Biochemistry and the
Protein Engineering Network of Centres of Excellence
University of Alberta

19970304 029

November 1996

WARNING

"The use of this information is permitted subject to recognition of proprietary and patent rights."

CRAD

National Defence Défense nationale

Canada

UNCLASSIFIED

UNCLASSIFIED

Suffield Report No. 638

High Resolution Liquid Chromatographic Mass Spectrometric and Tandem Mass Spectrometric Identification of Peptides using Electrospray Ionization

by

J.R. Hancock, P.A. D'Agostino and L.R. Provost

Defence Research Establishment Suffield

and

P.D. Semchuk and R.S. Hodges

Department of Biochemistry and the

Protein Engineering Network of Centres of Excellence

University of Alberta

UNCLASSIFIED

ABSTRACT

High resolution liquid chromatography electrospray mass spectrometry (LC-ESI-MS) was investigated for the identification of known and unknown peptides in a research effort designed to evaluate the applicability of this and complementary MS techniques for the characterization and identification of peptides. The monoisotopic molecular weights of five related peptides with molecular weights between 2000 and 2500 Da, typical of bioactive peptides, were acquired with a magnetic sector resolution of 3000 (10% valley definition). Tryptic maps were generated for each peptide during high resolution LC-ESI-MS analysis and collisionally activated dissociation (CAD) in the ESI interface resulted in the production of characteristic product ions that enabled amino acid sequencing of the tryptic fragments. Under narrow and wide mass range magnetic sector scanning conditions monoisotopic molecular weight errors were typically in the 10-20 and 30-40 ppm range, respectively. Unknown identification, a critical requirement for MS methods developed in support of chemical/biological sample analysis, was demonstrated during analysis of an incomplete synthetic peptide reaction mixture. The reaction mixture for the peptide, LTTAVKKVLTTGLPALIS, was selected since the 18 amino acid target peptide was not synthesized. In its place were a number of lower molecular weight unknown peptides. The monoisotopic molecular weights of these unknowns were determined to within 10-20 ppm with a magnetic sector resolution of 3500 (10% valley definition). However, co-elution of the unknown peptides during LC-ESI-MS analysis precluded the use of CAD-MS for amino acid sequencing. Amino acid sequence information for six unknown peptides was generated by ESI-MS/MS, making use of the quadrupole collisional cell associated with the hybrid tandem mass spectrometer. The product spectra acquired for the $(M+H)^+$ precursor ions for these peptides were dominated by b_n-series ions from which the amino acid sequence could be assigned. Supplementary information on the amino acid sequence was provided by the presence of y_n-series and internal fragment ions.

Executive Summary

Title: J.R. Hancock, P.A. D'Agostino, L.R. Provost, P.D. Semchuk and R.S. Hodges " High Resolution Liquid Chromatographic Mass Spectrometric and Tandem Mass Spectrometric Identification of Peptides using Electrospray Ionization", Suffield Report No. 638, 1996, UNCLASSIFIED.

Introduction: The Canadian Forces (CF) may be called on to conduct peacekeeping or peacemaking operations in regions of the world where there is a significant threat of chemical/biological warfare agent use. To operate effectively in these theatres the CF must be able to identify the exact nature of the chemical/biological agent(s). Mass spectrometry (MS) is a powerful analytical technique for the identification of both known and unknown compounds and DRE Suffield is currently investigating the applicability of this instrumental technique to fulfil CF agent detection and identification requirements.

Results: High resolution liquid chromatography electrospray mass spectrometry (LC-ESI-MS) was investigated for the identification of known and unknown peptides in a research effort designed to evaluate the applicability of this and complementary MS techniques for the characterization and identification of peptides. The monoisotopic molecular weights of five related peptides with molecular weights between 2000 and 2500 Da, typical of bioactive peptides, were acquired with a magnetic sector resolution of 3000 (10% valley definition). Tryptic maps were generated for each peptide during high resolution LC-ESI-MS analysis and collisionally activated dissociation (CAD) in the ESI interface resulted in the production of characteristic product ions that enabled amino acid sequencing of the tryptic fragments. Under narrow mass range magnetic sector scanning conditions monoisotopic molecular weight errors were typically in the 10-20 ppm range. Unknown identification, a critical requirement for MS methods developed in support of chemical/biological sample analysis, was demonstrated during analysis of an incomplete synthetic peptide reaction mixture. The target peptide was not synthesized but in its place were a number of lower molecular weight unknown peptides. The monoisotopic molecular weights of these unknowns was determined to within 10-20 ppm with a magnetic sector resolution of 3500. However, co-elution of the unknown peptides during LC-ESI-MS analysis precluded the use of CAD-MS for amino acid sequencing. Amino acid sequence information for six unknown peptides was generated by ESI-MS/MS, making use of the quadrupole collisional cell associated with the hybrid tandem mass spectrometer. The product spectra acquired for the $(M+H)^+$ precursor ions for these peptides were dominated by b_n-series ions from which the amino acid sequence could be assigned. Additional information was provide by y_n-series and internal fragment ions.

Significance of Results: The CF may be deployed in regions of the world where there is a significant threat of chemical/biological warfare agent use. Identification of the agent is of importance since the results of such analyses would contribute to the development of strategic and political positions regarding future Canadian military operations and would facilitate the dissemination of technical advice to in-theatre field commanders and medical personnel.

Future Goals: The CB threat spectrum includes chemical and biological warfare agents and toxins of biological origin in the "mid-spectrum" between these agents. The identification research effort has been focused on the detection and identification of these toxins of biological origin. Use of these warfare agents could easily go unconfirmed, as analytical methods have not been fully developed for their identification. DRE Suffield is now actively addressing this deficiency through the application and development of MS methods for identification of these agents.

TABLE OF CONTENTS

INTRODUCTION

The chemical/biological (CB) agent spectrum encompasses a wide range of toxic materials including chemical warfare agents, mid-spectrum agents (e.g., toxic peptides and proteins) and biological warfare agents. Recent advances in biotechnology, including the isolation and production of peptides using solid-phase synthesis and recombinant DNA-modified microorganisms, has opened up new avenues for the preparation of significant quantities of agents in the "mid-spectrum" between classical chemical and classical biological warfare agents. The current lack of appropriate analytical methods makes detection and identification of mid-spectrum agents extremely difficult. The Canadian Forces (CF) may be called on to perform peacekeeping or battlefield operations in regions of the world where there is a significant threat of chemical/biological warfare (CBW) agent use. To operate effectively in these theatres the CF must be able to identify the exact nature of the chemical/biological agent(s). This requirement is being actively addressed at DRES by conducting research and development into new analytical methods for the identification of chemical/biological warfare agents.

Mass spectrometry combines exceptional sensitivity, specificity and speed for the analysis of toxic peptides and proteins. Growth in this area has been extremely rapid, due to the recent development and commercial availability of two new ionization techniques, electrospray (ESI) and matrix-assisted laser desorption/ionization (MALDI) (1-3). The large number of potential toxins of biological origin precludes the development of individual methods for all the candidate toxins. Therefore, a general approach that provides maximal structural information has been adopted. Briefly, the strategy adopted for the identification of toxins at DRES using liquid chromatography (LC) and ESI-MS is as follows:

i) High resolution ESI-MS will be used to determine the monoisotopic or average molecular weight of target peptides or proteins as accurately as possible. This information may be

used to search peptide/protein databases for tentative identification purposes. With increasing molecular weight accuracy the number of possible database hits will be reduced.

ii) Amino acid sequencing of peptides will be performed through the use of collisionally activated dissociation (CAD)-MS or MS/MS with a hybrid tandem mass spectrometer. High or low resolution sequence ion masses may be accessed during CAD-MS or MS/MS, respectively. This information, when combined with accurate molecular weight data, could result in a single hit when searching peptide/protein databases.

iii) Higher molecular weight peptides or proteins require enzymatic or chemical degradation to produce low molecular weight peptide fragments that may be sequenced in the manner described above. The accurate molecular weight of the intact peptide/protein and its peptide fragments, referred to as a tryptic map, may be used to search peptide/protein databases. Amino acid sequence information for some of the peptide fragments provides additional information that would greatly reduce the number of database hits.

In the early 1980's Barber and co-workers revolutionized mass spectrometry by demonstrating the use of fast atom bombardment (FAB) mass spectrometry for the analysis of peptides (4). In the ensuing years, the useful mass range of mass spectrometry for biomolecule applications has increased, particularly with the demonstration by Fenn and co-workers that electrospray ionization could be used to form multiply charged gaseous ions from large biomolecules (5,6). Biemann (3) has reviewed the mass spectrometry of peptides and proteins and the current status of biological mass spectrometry was reviewed in 1994 by Burlingame, Boyd and Gaskell (7).

The electrospray interface was initially interfaced to a quadrupole mass spectrometer (1,2) and most of the applications in the literature deal with this type of instrumentation. Development of suitable electrospray interfaces for high resolution sector use (8) was in large part driven by the

potential to increase mass measurement accuracy. Use of high resolution enables the assignment of charge state to multiply charged isotope clusters and aids in the interpretation of peptide primary sequence data during collisionally activated dissociation (CAD) in the region between the capillary exit and skimmer in the electrospray interface (9-15). Starrett and DiDonato, working at a resolution of 5000 (5% valley) under voltage scanning conditions over a narrow mass range concentrated on the accurate mass measurement of product ions generated during CAD/MS (15). The value of high resolution for the assignment of charge state for CAD/MS product ions was demonstrated by Loo's group in a paper focusing on the determination of protein structural information following electrospray introduction (16).

Larsen and McEwen employed resolutions of 5000 and 10000 (10% valley) for accurate molecular weight determination and found that errors seldom exceeded 25 ppm for several pure peptides. Calibration was done internally and the isotopic cluster for the +5 charge state of insulin was resolved (9). Similar ppm errors were reported in a second paper for a number of standard peptides and proteins under lower resolution conditions during average molecular weight determinations. Average molecular weight accuracy was sufficient to allow for the differentiation of a single point modification differing by only 1 Da for biomolecules up to 20000 Da (17).

The molecular weights for a series of thirty-seven unknown synthetic peptides, used in research studies involving synthetic vaccines, antibacterial peptides or the de novo design of helical peptides and proteins, were determined with a magnetic sector instrument (18). All data were obtained with external calibration over a wide mass range during magnetic scanning. Errors between observed and theoretical monoisotopic molecular weights were typically in the 5 to 60 ppm range for the unknowns at sector resolutions between 2500 and 9000 (10% valley). Isotopic clusters for charge states up to +10 were resolved through the use of high resolution. CAD in the electrospray interface resulted in product ions that enabled either full or partial sequencing of most unknown peptides below 2000 Da.

The characterization and identification of peptides and proteins involves obtaining both molecular weight and amino acid sequence data from the intact peptide or peptide fragments formed following enzymatic or chemical degradation. Tryptic digestions, where the peptide or protein cleaves at the C-terminal of lysine or arginine amino acid residues, have been used frequently to produce peptide fragments that may be used to characterize the intact peptide/protein. Low resolution LC-MS and LC-MS/MS using fast atom bombardment (19), ion spray (20) and electrospray ionization (21) have been used frequently to characterize both intact peptides and their tryptic fragments. The molecular weights of the observed peptide fragments forms a tryptic map for that peptide or protein that further characterizes the intact peptide or proteins (19-24). The use of magnetic sector instruments for high resolution ESI-MS measurement of molecular or product ions is limited, with prior publications dealing with loop rather than chromatographic sample introduction (7). Chromatographic separations prior to high resolution ESI-MS characterization would be preferred at DRES since chemical/biological defence samples typically contain multiple sample components.

High resolution LC-ESI-MS was investigated for the identification of a number of known and unknown peptides in an effort designed to evaluate the applicability of this and complementary MS techniques for the characterization and identification of peptides. The monoisotopic molecular weights of five related peptides with molecular weights between 2000 and 2500 Da, typical of bioactive peptides, were acquired with a sector resolution of 3000 (10% valley definition). Tryptic maps were acquired for each peptide and high resolution LC-CAD-MS was investigated for the determination of tryptic fragment amino acid sequences. Unknown identification, a critical requirement for MS methods developed in support of chemical/biological sample analysis, was evaluated with a peptide reaction mixture. The reaction mixture for the incomplete synthesis of the peptide, LTTAVKKVLTTGLPALIS, was selected since the 18 amino acid target peptide was not synthesized. In its place were a number of lower molecular weight unknown peptides. Co-elution of the unknown peptides during LC-ESI-MS precluded the use of CAD-MS for amino acid

sequencing. ESI-MS/MS, making use of the quadrupole collisional cell, was therefore evaluated for the determination of amino acid sequences for the unknowns.

EXPERIMENTAL

Samples

Five related peptides with molecular weights between 2000 and 2500 Da and a sample of the reaction mixture from the incomplete synthesis of the peptide, LTTAVKKVLTTGLPALIS, were provided by Dr. R. Hodges' Group (Protein Engineering Network of Centres of Excellence, University of Alberta, Edmonton). The five peptides were shipped and received at Defense Research Establishment Suffield as frozen aqueous solutions (0.5 mL) in the 0.4 to 1.1 mg/mL range. The incomplete peptide reaction mixture was received as 1mg of a lyophilized powder and was dissolved in 1mL of distilled water. All samples were kept frozen at -18 °C until just prior to LC-ESI-MS analysis.

Tryptic digests were performed by dissolving modified sequencing grade trypsin (Boehringer Mannheim, Mannheim, Germany) in 0.1M ammonium bicarbonate and adding 50 μL of this trypsin solution (0.01 mg/mL) to 50 μL of the peptide, giving an enzyme:substrate ratios of between 1:40 to 1:110. Digestions were carried out for about 18 hours at 37 °C and were quenched with 10 μL of 10%TFA (to pH 2-3) prior to LC-ESI-MS analysis.

Instrumental

All electrospray mass spectra were acquired using a VG (Fisons) Autospec-Q hybrid tandem mass spectrometer (Manchester, UK) equipped with a VG (Fisons) Mark II electrospray interface.

The electrospray needle was operated at 7.6 kV and ions were accelerated into the mass spectrometer at 4 kV. Sampling cone voltages of 75 to 150 volts were utilized. Nitrogen (Very Dry, Liquid Carbonic Inc., Scarborough, Ont., Canada) bath gas was introduced into the interface (80 °C) at a flow rate of 500 L/hr. Nitrogen nebulizer gas was introduced at a flow rate of 14 L/hr. The electrospray interface was pumped with both a rotary and a turbomolecular pump, which enabled maintenance of a $3x10^{-6}$ and $5x10^{-8}$ Torr within the source and analyzer regions of the instrument, respectively.

LC-ESI-MS data for the five related peptides and their tryptic fragments were acquired in the continuum mode by scanning the magnet sector from 1400 to 400 Da (12 sec/decade) or 670 to 230 Da (15 sec/decade). Monoisotopic molecular weight data for the unknown components in the incomplete synthetic peptide mixture were acquired during LC-ESI-MS by magnetic sector scanning over 1900 to 400 Da (10 sec/decade) or 910 to 575 Da (30 sec/decade) in the continuum mode. Five to ten scans were typically averaged to enhance the signal-to-noise ratio and the data were smoothed using VG (Fisons) OPUS software. Resolutions of 2600 to 3500 (10% valley definition) were employed during magnetic sector scanning to facilitate accurate mass measurement of the ions. External calibrations were performed with solutions of polyethylene glycol (PEG) 200, 600 and 1540 (0.5 mg/mL in distilled water).

ESI-MS/MS data for the incomplete synthetic peptide were acquired following loop injection with a quadrupole CAD cell energy of 75 or 100 volts and an argon pressure of approximately 1×10^{-4} Torr near the CAD cell. The preferred conditions attenuated the precursor ions by about 60%-70% (CAD cell energy of 100 volts) and represented a good compromise between sensitivity and spectral content for these analyses. Product spectra were obtained at unit resolution by scanning the quadrupole from 800 to 80 Da (10 sec/scan). The magnetic sector resolution was set at 1000 (10% valley definition) for these analyses.

All peptide separations were performed with an Applied Biosystems Model 140B dual syringe pump (Foster City, CA) equipped with a 15 cm × 0.53 mm i.d. C_{18} (5 μm) packed J&W DB-1 fused-silica capillary column (Courtesy of Mr. L. Hogge and Mr. D. Olson, NRC, Saskatoon, Canada) and a Rheodyne 8125 (Cotati, CA) with a 5 μL sample loop. The following solvent compositions were prepared for sample introduction: Solvent A (0.05% TFA in water) and Solvent B (0.05% TFA in acetonitrile/water (80:20)). Chromatographic separations were performed using a 5% to 75%B gradient over 35 minutes. In order to minimize dead volume and ensure reproducible mixing, the mobile phase was delivered at 200 μL/min and split prior to the injector such that the flow through the column was 20 μL/min. Loop injections (20 μL) of the incomplete synthetic peptide mixture and calibration standards were made with a second Rheodyne 8125 located after the column under isocratic conditions (50% B, 10 μL/min).

RESULTS AND DISCUSSION

DRES has undertaken a research program focussed on the application of mass spectrometry (MS) for the unambiguous identification of toxic peptides and proteins in man-made and environmental matrices (25) in order to meet the Canadian Forces requirement of an independent chemical/biological agent detection and identification capability. Recently DRES analysed a series of synthetic peptides which were provided as unknowns by ESI-MS using loop injection (18). Monoisotopic molecular weight and partial primary sequence data were obtained for these peptides under high resolution conditions using a hybrid tandem instrument. With increasing molecular weight the amount of amino acid sequence information that could be accessed by CAD/MS conditions was reduced. Enzymatic digestions with reagents such as trypsin are commonly employed in order to cleave peptides and proteins into smaller peptide fragments that can be sequenced by CAD/MS. A series of five related peptides were selected from the original set of 37 peptides (18), digested with trypsin and the resulting tryptic fragments analysed by LC-ESI-MS. It is unlikely that

toxic peptides or proteins would be of laboratory purity when used in a biological weapons scenario. In all likelihood, they will be the product of an incomplete synthesis and will contain precursors and byproducts as well as the target peptide or protein. An incomplete synthetic peptide mixture was therefore obtained to simulate this scenario. The 18 amino acid peptide, LTTAVKKVLTTGLPALIS, was not synthesized and in its place were a number of other unknown peptides. This sample was analysed by LC-ESI-MS and ESI-MS/MS and six unknown peptides were identified on the basis of monoisotopic molecular weight and amino acid sequence data.

LC-ESI-MS Analysis of Five Related Peptides

Five related peptides (numbered 11, 12, 15, 16 and 18), with amino acid sequences listed in Table I, were selected to investigate the amount amino acid sequence information that could be accessed following tryptic digestion of each peptide. Figure 1 illustrates the LC-ESI-MS total-ion-current chromatogram obtained for this mixture of peptides with a magnetic sector resolution of 3000. All five peptides were resolved in under 30 minutes, with a gradient of 2%B/min. Isotopic clusters of the general form, $(M+nH)^{n+}$, were observed using a sampling cone voltage of 75 volts, while at a higher sampling cone voltages, 150 volts, charge stripping associated with the CAD process increased the presence of lower charged isotopic clusters (Figure 2). The spectra did not contain product ions indicative of the amino acid sequence, with the only product ions being due to loss of water from the $(M+3H)^{3+}$ ion.

The monoisotopic molecular weights of each of the peptides was calculated from the ^{12}C only ion in the isotopic clusters and the calculated values were compared to theoretical values. Table I lists both the theoretical and calculated monoisotopic molecular weights for the 5 peptides. Errors associated with the monoisotopic molecular weights when scanning from 1400 to 400 Da were in the 30-40 ppm range, sufficiently small to significantly reduce the number of possible hits when searching a peptide/protein database for possible matches and comparable to the 5 to 60 ppm range

found during loop as opposed to chromatographic sample introduction (18).

In order to access amino acid sequence information, the peptides were each enzymatically digested using trypsin and the digests analysed by LC-ESI-MS with a magnetic sector resolution of 3000. Figure 3 illustrates reconstructed-ion-chromatograms obtained during the analysis of peptide 18 and clearly illustrates the presence of six components, whereas only five fragments were predicted from the original structure of the peptide. Theoretical tryptic maps based on cleavage on the C-side of lysine were generated for all five peptides and all the expected tryptic fragments were detected with the exception of EQEK for peptide 12. These data, listed in Table II, were acquired over a narrower mass range (670 to 230 Da) and errors between observed and theoretical monoisotopic molecular weight were typically in the 10-20 ppm range. Several tryptic fragments that had not undergone complete cleavage at the C-side of every lysine were also observed, as well as evidence of deamidation at the C-terminal.

The similarity in peptide structure resulted in the formation of a number of common tryptic fragments. Higher voltage sampling cone conditions resulted in the acquisition of CAD/MS data that contained product ions indicative of the amino acid sequence of these fragments. Figure 4a illustrates the CAD/MS data obtained for tryptic fragments ELEK with a sampling cone voltage of 150 volts. Product ions in the y_n-series, corresponding to cleavage of E (y_3) and L (y_2) amino acid residues from the $(M+H)^+$ ion were observed at m/z 389.2413 and m/z 276.1521, respectively. Leucine cannot be differentiated from isoleucine during CAD/MS analysis and in all cases the isomer that would be present in the tryptic fragment has been assumed in the structures presented. The mass of the remaining amino acid residues would be 257.1337 Da (based on the y_2 ion at m/z 276.1521) or 257.1426 (based on the b_2 ion at m/z 243.1272). Possible two amino acid residue combination with a nominal mass of 257 Da are E and Q (257.1012 Da), W and A (257.1164 Da), E and K (257.1376 Da) and R and T (257.1488 Da). The last two combinations fall within the errors typically recorded at his resolution (Table III). However an intense loss of H_2O from the y_2 ion would be favoured for

a tryptic fragment with an N-terminal E (giving rise to pEK where E cyclizes to pE, pyroglutamic acid). This evidence would suggest an amino acid sequence of ELEK.

The tryptic fragment, LLK, exhibited a single y_2 product ion at m/z 260.1940, which indicated the presence of an N-terminal leucine (Figure 4b and Table III). The observed amino acid residue mass for the remainder of the peptide was 241.1756 based on the y_2 ion (m/z 260.1940) Da. Two possible two amino acid residue combinations, Q and L (241.1426 Da) and K and L (241.1790 Da), have a nominal mass of 241 Da. In this case the amino acid sequence could be assigned as LLK based on the differences between observed and theoretical masses.

The tryptic fragment from the N-terminal, Ac-ELEK, exhibited y_n-series ions at m/z 389.2422 (y_3) and m/z 276.1500 (y_2) due to cleavage of Ac-E and L, respectively (Figure 5a). The mass of the remaining amino acid residues would be 257.1316 Da (based on the y_2 ion at m/z 276.1500) or 257.1399 (based on the b_2 ion at m/z 285.1379). Possible two amino acid residue combination with a nominal mass of 257 Da are E and Q (257.1012 Da), W and A (257.1164 Da), E and K (257.1376 Da) and R and T (257.1488 Da). The last two combinations fall within the errors typically recorded at his resolution (Table III). However once again an intense loss of H_2O from the y_2 ion would be favoured for a tryptic fragment with an N-terminal E (giving rise to pEK). This evidence would suggest an amino acid sequence of Ac-ELEK.

Based on molecular weight, an unexpected tryptic fragment for peptide 18 was believed to have an amino acid sequence of LLKEWEK or EWEKLLK, the result of incomplete cleavage of a peptide bond on the C-side of lysine. Figure 5b illustrates the CAD/MS data obtained for this tryptic fragment. A y_6^{2+} ion at 416.7337 indicated loss of L from the N-terminal and suggested LLKEWEK as opposed to EWEKLLK. This was further supported by the y_4 and y_3 ions at m/z 591.2747 and 462.2397, respectively (Table III).

Analysis of An Incomplete Synthetic Peptide Reaction Mixture

Samples suspected of containing toxic peptides and proteins that are analysed at DRES will originate from various sources. Samples collected during CF operations will likely contain multiple components including those from incomplete synthesis. Analytical methods will be required to establish the identity of the principal sample components under this likely scenario. An incomplete synthetic peptide reaction mixture for, LTTAVKKVLTTGLPALIS, was obtained to evaluate the amount of information that could be extracted during mass spectrometric analysis.

Preliminary LC-ESI-MS of the sample indicated that the synthesis failed to produce the desired product, but a number of other unknown peptides, presumably related to the desired peptide product were observed. Figure 6 illustrates the LC-ESI-MS total-ion-current chromatogram obtained under a fairly rapid gradient of 2%B/min. Most of the total-ion-current was due to six peptides, with three of the components co-eluting under the LC conditions employed. They were however resolved by mass under lower sampling cone voltage conditions (75 volts) that do not favour product ion production (Figure 6 inset). Monoisotopic molecular weights were calculated for each of the six unknown peptides and compared to theoretical data following ESI-MS/MS amino acid sequencing of the peptides (Table IV). Monoisotopic molecular weight errors under wide and narrow mass range magnetic scanning conditions were in the 35 to 60 ppm and 5 to 20 ppm range, respectively. It was not possible to distinguish between leucine and isoleucine and assignment of these amino acid residues was based solely on the amino acid sequence of the desired product.

Electrospray ionization typically produces multiply protonated molecular ions and under CAD/MS conditions frequently produces y_n-, b_n- and a_n- series product ions that aid in the determination of amino acid sequence. The observed peptides yielded product ions from these series as well as a number of internal fragment ions. The formation of internal fragments requires the cleavage of two peptide bonds and charge retention by the resulting fragment ion (26). In order to

describe the internal fragment ion, the nomenclature suggested by Roepstorff and Fohlman (27) was used with slight modification. The general form for internal fragment ions is $(b_r y_s)_{(r+s-t)}$ where r indicates the bond cleaved counting from the N-terminal, s indicates the bond cleaved counting from the C-terminal and t indicated the total number of amino acid residues in the peptide. The subscript (r + s - t) therefore indicates the total number of amino acid residues in the internal fragment. For example, a peptide with the sequence LPALIS with an internal fragment PAL, has r=4, s=5, t=6 would be described as an $(b_4y_5)_3$ ion. The presence of internal fragments can be useful in reconstructing internal amino acid sequence for unknown peptides.

In general CAD/MS amino acid sequence information is preferred over MS/MS data as CAD/MS data may be acquired under high resolution conditions. Data collected under MS/MS condition using the quadrupole mass analyser is typically limited to unit resolution, making resolution of multiply charged isotopic clusters more difficult (or impossible) and assignment of m/z less accurate. The major advantage of MS/MS lies in the its ability to differentiate sample components on the basis of mass, thereby reducing the need for chromatographic separation. In the case of the incomplete synthetic peptide mixture, chromatographic resolution of all the sample components in a reasonable timeframe would be unlikely. As a result CAD/MS data would contain product and precursor ions for all co-eluting sample components, thus complicating the interpretation of the observed product ions. ESI-MS/MS was therefore used for the determination of amino acid sequence since this technique allows mass selection of the precursor ion prior to CAD in the quadrupole cell.

Figure 7a illustrates the product spectrum of m/z 613.4, the $(M+H)^+$ ion of the first unknown peptide. The spectrum contained an intense series of b_5 to b_2 ions, their corresponding a_n-series ions and y5, y3 and y2 ions (Table V). The presence of these series was sufficient to fully sequence this peptide: LPALIS. The product spectrum of m/z 670.4, the $(M+H)^+$ ion for the second peptide, was dominated by b_6 to b_2 series ions. Several a_n- and y_n-series ions were also detected but neither a b_1

or a y_6 ion, necessary to establish the N-terminal was observed (Table V). Two possibilities exist that would account for a two amino acid residue mass of 170 Da. G and I/L or V and A would both be possible with GL being more probable given the fact that the desired peptide contained GLPALIS on the C-terminal. Internal fragment ions, with the most intense being at m/z 282 and m/z 169, provided complementary internal sequence information.

The presence of strong $(M+H)^+$ ions is also observed in the product ion spectra for the peptides AGLPALIS and TGLPALIS which are illustrated in Figure 8. Again the spectra are dominated by b-series ions which allow most of the primary sequence of these peptides to be determined. Significant internal fragment ions were observed for both peptide, primarily related to b series ions. Figure 9 illustrates the product ion spectra obtained for the peptides VGLPALIS and LGLPALIS. The spectra exhibit strong $(M+H)^+$ ions and are dominated by b-series ions which provide almost full sequence data for these peptides. Table V lists the product ions and internal fragment ions observed under ESI-MS/MS conditions for the six peptides identified in the incomplete synthetic reaction mixture. In general the product ion spectra contained sufficient b-series ions that most of the primary sequence for the peptides could be assigned. In addition to b-series ions the corresponding a-series ions (at a mass 28 Da less than the corresponding b-series ion) were also observed for most of the product ions. Few y-series ions were observed, but a significant y_5 ion at m/z 500 was present for all six peptides. Numerous internal fragments were observed with the most intense being related to the fragments PA, PAL and PALI.

CONCLUSIONS

High resolution liquid chromatography electrospray mass spectrometry (LC-ESI-MS) was investigated for the identification of known and unknown peptides in a research effort designed to evaluate the applicability of this and complementary MS techniques for the characterization and identification of peptides. The monoisotopic molecular weights of five related peptides with molecular weights between 2000 and 2500 Da, typical of bioactive peptides, were acquired with a magnetic sector resolution of 3000 (10% valley definition). When no significant CAD ions were observed, the samples were enzymatically digested using trypsin to produce smaller peptide fragments suitable for amino acid sequencing. Tryptic maps were generated for each peptide during high resolution LC-ESI-MS analysis and collisionally activated dissociation in the ESI interface resulted in the production of characteristic product ions that enabled amino acid sequencing of the tryptic fragments. Under narrow and wide mass range magnetic sector scanning conditions monoisotopic molecular weight errors were typically in the 10-20 and 30-40 ppm range, respectively.

Unknown identification, a critical requirement for MS methods developed in support of chemical/biological sample analysis, was demonstrated during analysis of an incomplete synthetic peptide reaction mixture. The reaction mixture for the peptide, LTTAVKKVLTTGLPALIS, was selected since the 18 amino acid target peptide was not synthesized. In its place were a number of lower molecular weight unknown peptides. The monoisotopic molecular weights of these unknowns were determined to within 10-20 ppm with a magnetic sector resolution of 3500 (10% valley definition). However, co-elution of the unknown peptides during LC-ESI-MS analysis precluded the use of CAD-MS for amino acid sequencing. Amino acid sequence information for six unknown peptides was generated by ESI-MS/MS, making use of the quadrupole collisional cell associated with the hybrid tandem mass spectrometer. The product spectra acquired for the $(M+H)^+$ precursor ions for these peptides were dominated by b_n-series ions from which the amino acid sequence could be assigned. Supplementary information on the amino acid sequence was provided by the presence of y_n-series and internal fragment ions.

REFERENCES

1. Meng, C.K., Mann, M. and Fenn, J.B., Z. Phys. D, **10**, 361, 1988.

2. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M., Science, **246**, 64, 1989.

3. Biemann, K., Annu. Rev. Biochem., **61**, 977, 1992.

4. Barber, M., Bordoli, R.S., Sedgwick, R.D. and Tyler, A.N., J. Chem. Soc. Chem. Commun., 325, 1981.

5 Meng, C.K., Mann, M. and Fenn, J.B., Z. Phys. D, **10**, 361, 1988.

6. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M., Science, **246**, 64, 1989.

7. Burlingame, A.L., Boyd, R.K. and Gaskell, S.J., Anal. Chem., **66**, 634R, 1994.

8. Meng, C.K., McEwen, C.N. and Larsen, B.S., Rapid Commun. Mass Spectrom., **4**, 147, 1990.

9. Meng, C.K., McEwen, C.N. and Larsen, B.S., Rapid Commun. Mass Spectrom., **4**, 151, 1990.

10. Larsen, B.S. and McEwen, C.N., J. Am. Soc. Mass Spectrom., **2**, 205, 1991.

11. Chapman, J.R., Gallagher, R.T., Barton, E.C., Curtis, J.M. and Derrick, P.J., Org. Mass Spectrom., **27**, 195, 1992.

12. McEwen, C.N. and Larsen, B.S., Rapid Commun. Mass Spectrom., **6**, 173, 1992.

13. Cody, R.B., Tamura, J. and Musselman, B.D., Anal. Chem., **64**, 1561, 1992.

14. Loo, J.A., Ogorzalek Loo, R.R. and Andrews, P.C., Org. Mass Spectrom., **28**, 1640, 1993.

15. Starrett, A.M. and DiDonato, G.C., Rapid Commun. Mass Spectrom., **7**, 12, 1993.

16. Loo, J.A., Ogorzalek Loo, R.R. and Andrews, P.C., Org. Mass Spectrom., **28**, 1640, 1993.

17. McEwen, C.N. and Larsen, B.S., Rapid Commun. Mass Spectrom., **6**, 173, 1992.

18. D'Agostino, P.A., Hancock, J.R., Provost, L.R., Semchuk, P.D. and Hodges, R.S., Rapid Commun. Mass Spectrom., **9**, 597, 1995.

19. Whaley, B. and Caprioli R.M., Bio. Mass Spectrom., **20**, 210, 1991.

20. Chang, J.P., Richardson, J.M. and Riggin, R.M., J. Liq. Chromatogr., **17**, 2881, 1994.

21. Hail, M., Lewis, S., Jardine, I., Liu, J. and Novotny, M., J. Microcol. Sep., **2**, 285, 1990.

22. Covey, T., Huang, E.C. and Henion J.D., Anal. Chem., **63**, 1193, 1991.

23. Ling, V., Guzzetta, A.W., Canova-Davis, E., Stults, J.T., Hancock, W.S., Covey, T.R. and Shushan, B.I., Anal. Chem., **63**, 2909, 1991.

24. Kassel, D.B., Musselman, B.D. and Smith, J.A., Anal. Chem., **63**, 1091, 1991.

25. Hancock, J.R., D'Agostino, P.A. and Provost, L.R., "Identification of Ile-Ser-Bradykinin in Aqueous samples by Liquid Chromatography with Ultraviolet and Electrospray Mass Spectrometric Detection", Suffield Memorandum No. 1465, June 1995, UNCLASSIFIED.

26. Ballard, K.D. and Gaskell S.J., Int. J. Mass Spectrom. Ion Proc., **111**, 173, 1991.

27. P.Roepstorff and J. Fohlman, Biomed. Mass Spectrom., 11 (1984) 601.

Table I: LC-ESI-MS analysis of synthetic peptides with a resolution of 3000 (10% valley definition) and a mass range of 1400 to 400 Da.

No.	Peptide Sequence	Theoretical Monoisotopic Mol. Wt.	Calculated Monoisotopic Mol. Wt.[a]	Error (ppm)
11	Ac-ELEKLLKELEKLLKEKEK-NH_2	2280.3668	2280.30±0.02	29
12	Ac-ELEKLLKEQEKLLKELEK-NH_2	2266.3511	2266.27±0.05	36
15	Ac-ELEKLLKECEKLLKELEK-NH_2	2255.2810	2255.21±0.03	31
16	Ac-ELEKLLKEYEKLLKELEK-NH_2	2315.3351	2315.26±0.02	32
18	Ac-ELEKLLKEWEKLLKELEK-NH_2	2238.3511	2338.28±0.02	32

[a] Monoisotopic molecular weights represent the average ± SD of four LC-ESI-MS analyses. During each analysis the average monoisotopic molecular weight was calculated using $(M+nH)^{n+}$ (n=2, 3 and 4 for a sampling cone voltage of 75 volts; n=2 and 3 for a sampling cone voltage of 150 volts) isotopic clusters.

Table II: Tryptic fragments observed during LC-ESI-MS analysis of tryptic digests of peptides 11, 12, 15, 16 and 18.

Tryptic Fragment (Theoretical Monoisotopic Mol. Wt.)	Observed Monoisotopic Molecular Weight [Resolution: 3000 (10% valley definition), Mass range 670 to 230 Da] (ppm error)				
	Peptide # 11	Peptide # 12	Peptide # 15	Peptide # 16	Peptide # 18
Ac-ELEK 559.2853	559.2882 (5)	559.2883 (5)	559.2806 (8)	559.2903 (9)	559.2903 (9)
LLK 372.2736	372.2735 (0.3)	372.2755 (5)	372.2691 (12)	372.2693 (12)	372.2710 (7)
ELEK 517.2748	517.2701 (9)	517.2747 (0.2)	517.2657 (18)	517.2686 (12)	517.2762 (3)
EK 275.1481	275.1508 (10)				
EK-NH_2 274.1641	274.1724 (30)				
EQEK 532.2493		not detected			
ELEK-NH_2 516.2908		516.2869 (8)	516.2780 (25)	516.2896 (2)	516.2886 (4)
ECEK 507.1999			507.1886 (22)		
EYEK 567.2540				567.2470 (12)	
EWEK 590.2700					590.2752 (9)

Table III: Sequence information for common tryptic fragments following LC-ESI-MS analysis with a resolution of 3000 (10% valley definition) and a sampling cone voltage of 150 volts.

Tryptic Fragment	Ion Identity	Theoretical Mass (Da)	Observed Mass (Da)	Error (Da)
a) ELEK	$(M+H)^+$	518.2826	518.2804	0.0022
	y_3	389.2400	389.2413	0.0013
	y_2	276.1559	276.1521	0.0038
	y_2-H_2O	258.1454	258.1462	0.0008
	b_2	243.1345	243.1272	0.0073
b) LLK	$(M+H)^+$	373.2815	373.2792	0.0023
	y_2	260.1974	260.1940	0.0034
c) Ac-ELEK	$(M+H)^+$	560.2932	560.2884	0.0048
	y_3	389.2400	389.2422	0.0022
	y_2	276.1559	276.1500	0.0059
	y_2-H_2O	258.1454	258.1391	0.0063
	b_2	285.1451	285.1379	0.0072
d) LLKEWEK	$(M+2H)^{2+}$	473.2744	473.2802	0.0058
	y_6^{2+}	416.7323	416.7337	0.0014
	y_4	591.2778	591.2747	0.0031
	y_3	462.2353	462.2397	0.0044

Table IV: LC-ESI-MS determination of monoisotopic molecular weights for peptides in an incomplete synthetic reaction mixture.

Sample Component (Peak No.)	Theoretical Monoisotopic Molecular Weight	Observed Monoisotopic Molecular Weight	
		Resolution: 3500 (10% valley definition), Mass Range: 910 to 575 Da (ppm error)	Resolution: 2600 (10% valley definition), Mass Range: 1900 to 400 Da (ppm error)
LPALIS (1)	613.3925	613.3883 (7)	613.3703 (36)
GLPALIS (2)	670.4139	670.4057 (12)	670.3854 (43)
AGLPALIS (3)	741.4510	741.4433 (10)	741.4231 (38)
TGLPALIS (4)	771.4616	771.4495 (16)	771.4194 (55)
VGLPALIS (5)	769.4823	769.4708 (15)	769.4360 (60)
LGLPALIS (6)	783.4980	783.4841 (18)	783.4575 (52)
	Average ppm error	13 ± 4 (n=6)	47 ± 10 (n=6)

Table V: ESI-MS/MS product ions for the peptides in incomplete synthetic reaction mixture.

Possible Ion Identity	Peptide Amino Acid Sequence m/z (% relative intensity)					
Series Ions	LPALIS	GLPALIS	AGLPALIS	TGLPALIS	VGLPALIS	LGLPALIS
b_7			636 (8)	666 (11)	664 (18)	678 (4)
b_6		565 (18)	523 (13)	553 (18)	551 (13)	565 (30)
b_5	508 (5)	452 (44)	410 (12)	440 (16)	438 (11)	452 (18)
b_4	395 (21)	339 (51)	339 (6)	369 (5)	367 (5)	381 (2)
b_3	282 (75)	268 (14)	242 (8)	272 (10)	270 (6)	284 (5)
b_2	211 (32)	171 (13)	129 (2)	159 (3)		171 (3)
a_7					636 (1)	650 (1)
a_6		537 (5)	495 (5)	525 (3)	523 (4)	537 (3)
a_5	480 (3)	424 (12)	382 (5)	412 (12)	410 (11)	424 (5)
a_4	367 (13)	311 (8)	311 (1)		339 (4)	353 (3)
a_3	254 (3)			244 (2)		
a_2	183 (13)	143 (11)				
y_6			613 (1)			613 (2)
y_5	500 (4)	500 (45)	500 (30)	500 (27)	500 (34)	500 (22)
y_3	332 (3)	332 (11)	332 (1)			
y_2	219 (6)	219 (9)	219 (1)			

Possible Ion Identity	Peptide Amino Acid Sequence m/z (% relative intensity)					
Internal Fragments	LPALIS	GLPALIS	AGLPALIS	TGLPALIS	VGLPALIS	LGLPALIS
GLPALI			565 (2) $(b_7y_7)_6$		565 (1) $(b_7y_7)_6$	565 (30) $(b_7y_7)_6$
GLPAL			452 (4) $(b_6y_7)_5$	452 (9) $(b_6y_7)_5$	452 (16) $(b_6y_7)_5$	452 (18) $(b_6y_7)_5$
PALI or LPAL	395 (21) $(b_5y_5)_4$	395 (11) $(b_6y_5)_4/(b_5y_6)_4$	395 (8) $(b_7y_5)_4/(b_6y_6)_4$	395 (13) $(b_7y_5)_4/(b_6y_6)_4$	395 (4) $(b_7y_5)_4/(b_6y_6)_4$	395 (14) $(b_7y_5)_4/(b_6y_6)_4$
LPAL-CO or PALI-CO	367 (13) $(a_5y_5)_4$	367 (9) $(a_6y_5)_4/(a_5y_6)_4$	367 (5) $(a_7y_5)_4/(a_6y_6)_4$	367 (5) $(a_7y_5)_4/(a_6y_6)_4$	367 (5) $(a_7y_5)_4/(a_6y_6)_4$	367 (7) $(a_7y_5)_4/(a_6y_6)_4$
GLPA			339 (6) $(b_5y_7)_4$		339 (4) $(b_5y_7)_4$	339 (21) $(b_5y_7)_4$
ALI	298 (2) $(b_5y_4)_3$	298 (3) $(b_6y_4)_3$		298 (9) $(b_7y_4)_3$	298 (4) $(b_7y_4)_3$	
PAL or LPA	282 (75) $(b_4y_5)_3$	282 (63) $(b_5y_5)_3/(b_4y_6)_3$	282 (36) $(b_6y_5)_3/(b_5y_6)_3$	282 (50) $(b_6y_5)_3/(b_5y_6)_3$	282 (35) $(b_6y_5)_3/(b_5y_6)_3$	282 (47) $(b_6y_5)_3/(b_5y_6)_3$
PAL-CO or LPA-CO	254 (3) $(a_4y_5)_3$	254 (4) $(a_5y_5)_3/(a_4y_6)_3$	254 (5) $(a_6y_5)_3/(a_5y_6)_3$	254 (3) $(a_6y_5)_3/(a_5y_6)_3$	254 (3) $(a_6y_5)_3/(a_5y_6)_3$	254 (1) $(a_6y_5)_3/(a_5y_6)_5$
LI	227 (7) $(b_5y_3)_2$	227 (11) $(b_6y_3)_2$	227 (4) $(b_7y_3)_2$	227 (18) $(b_7y_3)_2$	227 (5) $(b_7y_3)_2$	227 (7) $(b_7y_3)_2$
LP	211 (32) $(b_2y_6)_2$	211 (5) $(b_3y_6)_2$	211 (1) $(b_4y_6)_2$			
LI-CO	199 (8) $(a_5y_3)_2$	199 (6) $(a_6y_3)_2$	199 (7) $(a_7y_3)_2$	199 (5) $(a_7y_3)_2$		199 (2) $(a_7y_3)_2$
AL	185 (3) $(b_4y_4)_2$	185 (9) $(b_5y_6)_2$	185 (9) $(b_6y_4)_2$		185 (5) $(b_6y_4)_2$	185 (4) $(b_4y_4)_2$
PA	169 (34) $(b_3y_5)_2$	169 (55) $(b_4y_5)_2$	169 (25) $(b_5y_5)_2$	169 (23) $(b_5y_5)_2$	169 (14) $(b_5y_5)_2$	169 (20) $(b_5y_5)_2$
PA-CO	141 (8) $(a_3y_5)_2$	141 (29) $(a_4y_5)_2$	141 (3) $(a_5y_5)_2$	141 (4) $(a_5y_5)_2$	141 (5) $(a_5y_5)_2$	141 (4) $(a_3y_5)_2$

Figure 1: LC-ESI-MS total-ion-current (1400 to 400 Da) chromatogram for a mixture of synthetic peptides 11, 12, 15, 16 and 18 with a resolution of 3000 (10% valley definition) and a sampling cone voltage of 75 volts.

Figure 2: ESI-MS data for synthetic peptide 16 obtained during LC-ESI-MS analysis with a resolution of 3000 (10% valley definition) and a sampling cone voltage of a) 75 volts and b) 150 volts. (Insert: expansion of $(M+4H)^{4+}$ isotopic cluster).

Figure 3: Reconstructed-ion-current chromatograms for the $(M+H)^{+}$ ions of tryptic fragments of synthetic peptide 18 obtained during LC-ESI-MS analysis with a sector resolution of 3000 (10% valley definition) and a sampling cone voltage of 75 volts.

Figure 4: Typical CAD/MS data obtained for tryptic fragments a) ELEK and b) LLK obtained during LC-ESI-MS analysis with a sector resolution of 3000 (10% valley definition) and a sampling cone voltage of 150 volts.

Figure 5: Typical CAD/MS data obtained for tryptic fragments a) Ac-ELEK and b) LLKEWEK obtained during LC-ESI-MS analysis with a sector resolution of 3000 (10% valley definition) and sampling cone voltages of 150 and 115 volts, respectively.

Figure 6: LC-ESI-MS total-ion-current (910 to 575 Da) chromatogram of the peptides in an incomplete synthetic reaction mixture obtained with a resolution of 3500 (10% valley definition) and a sampling cone voltage of 75 volts. (Inset: $(M+H)^{+}$ isotopic clusters for peptides 2, 3 and 4).

Figure 7: Product ion mass spectra for peptides a) LPALIS and b) GLPALIS obtained during loop injection of the incomplete synthetic reaction mixture. (Quadrupole CAD cell: 100 volts, 1.1×10^{-4} Torr argon).

Figure 8: Product ion mass spectra for peptides a) AGLPALIS and b) TGLPALIS obtained during loop injection of the incomplete synthetic reaction mixture. (Quadrupole CAD cell: 100 volts, 1.1×10^{-4} Torr argon).

Figure 9: Product ion mass spectra for peptides a) VGLPALIS and b) LGLPALIS obtained during loop injection of the incomplete synthetic reaction mixture. (Quadrupole CAD cell: 100 volts, 1.1×10^{-4} Torr argon).

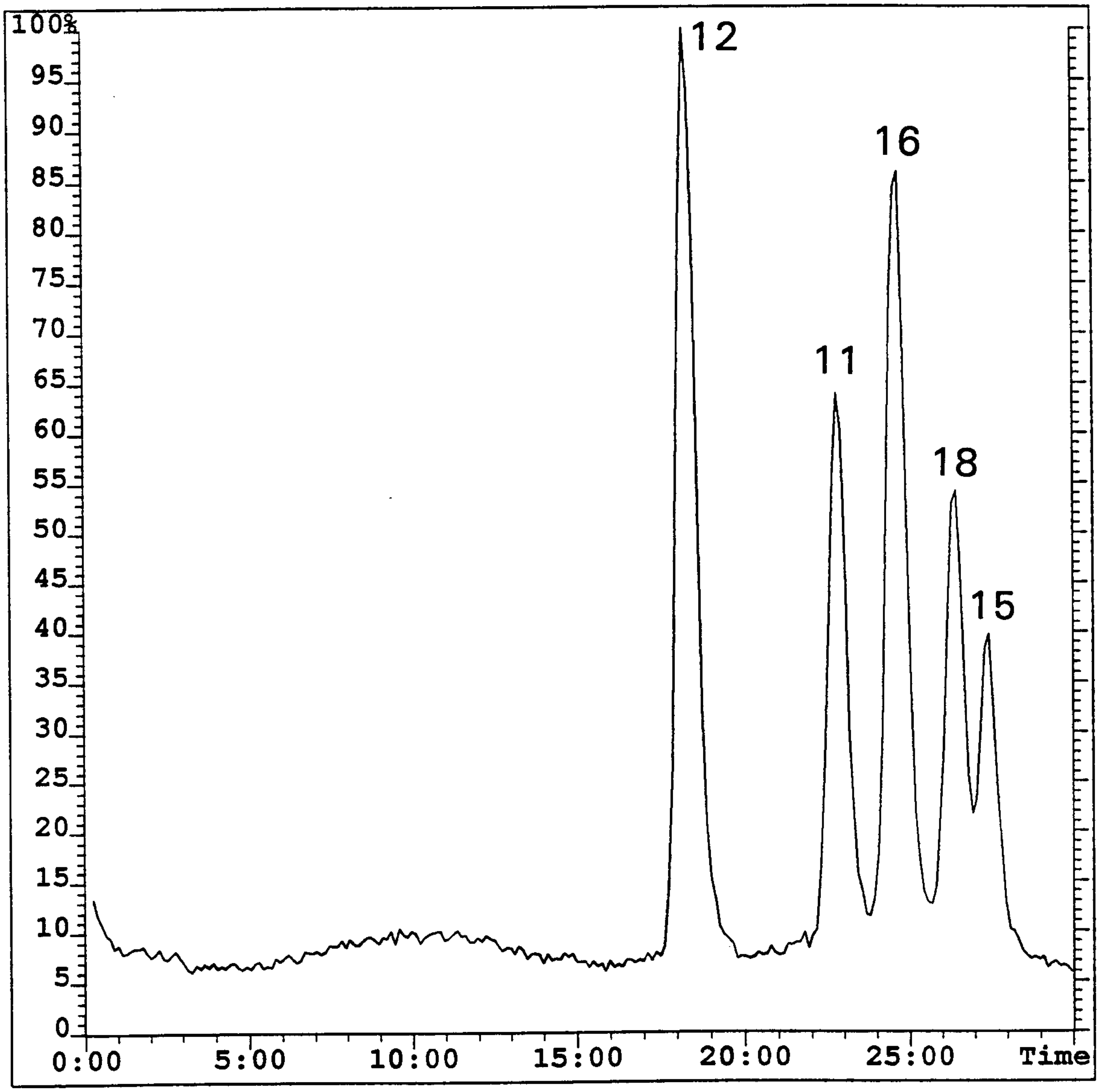

Figure 1: LC-ESI-MS total-ion-current (1400 to 400 Da) chromatogram for a mixture of synthetic peptides 11, 12, 15, 16 and 18 with a resolution of 3000 (10% valley definition) and a sampling cone voltage of 75 volts.

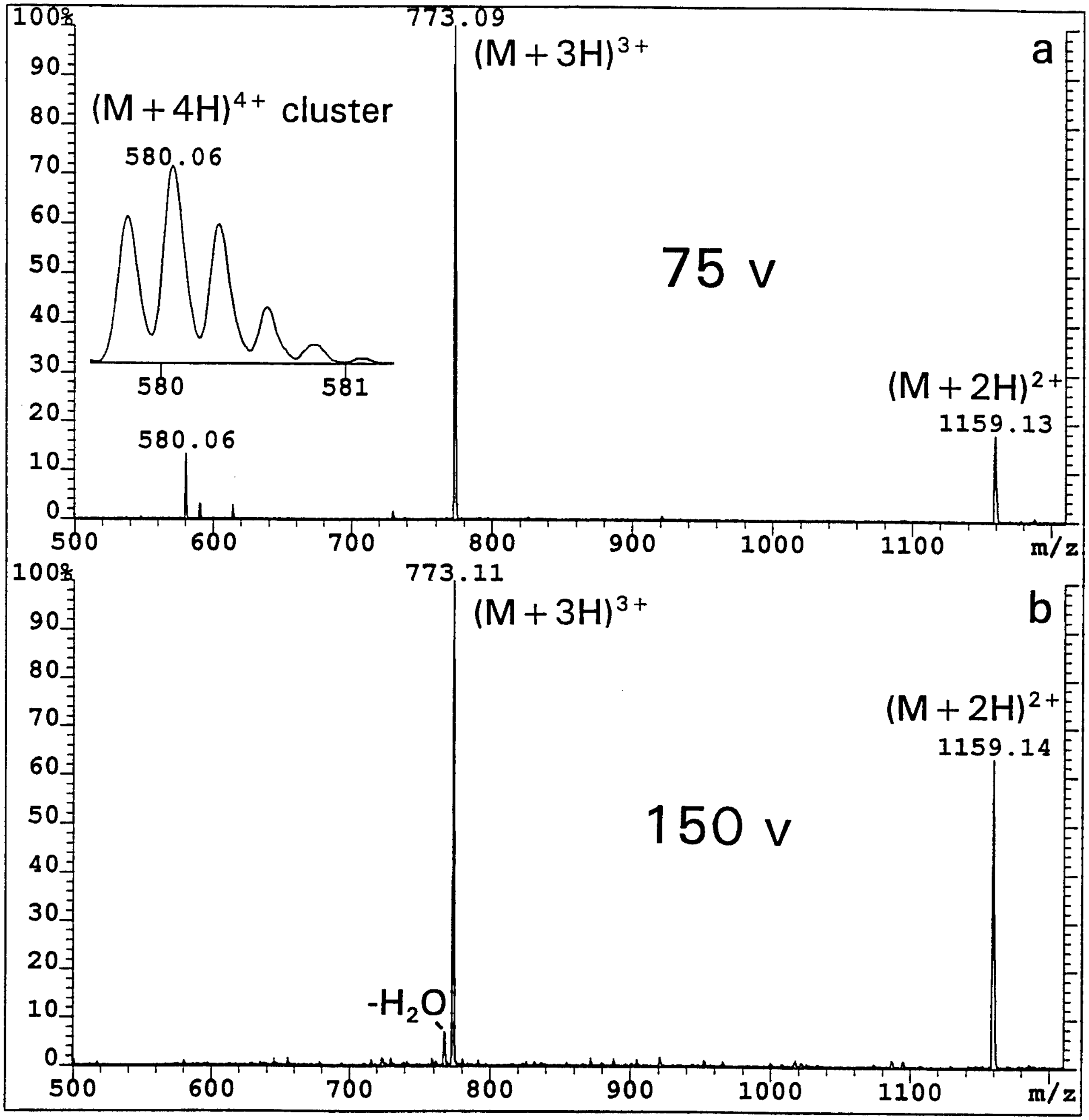

Figure 2: ESI-MS data for synthetic peptide 16 obtained during LC-ESI-MS analysis with a resolution of 3000 (10% valley definition) and a sampling cone voltage of a) 75 volts and b) 150 volts. (Insert: expansion of $(M+4H)^{4+}$ isotopic cluster).

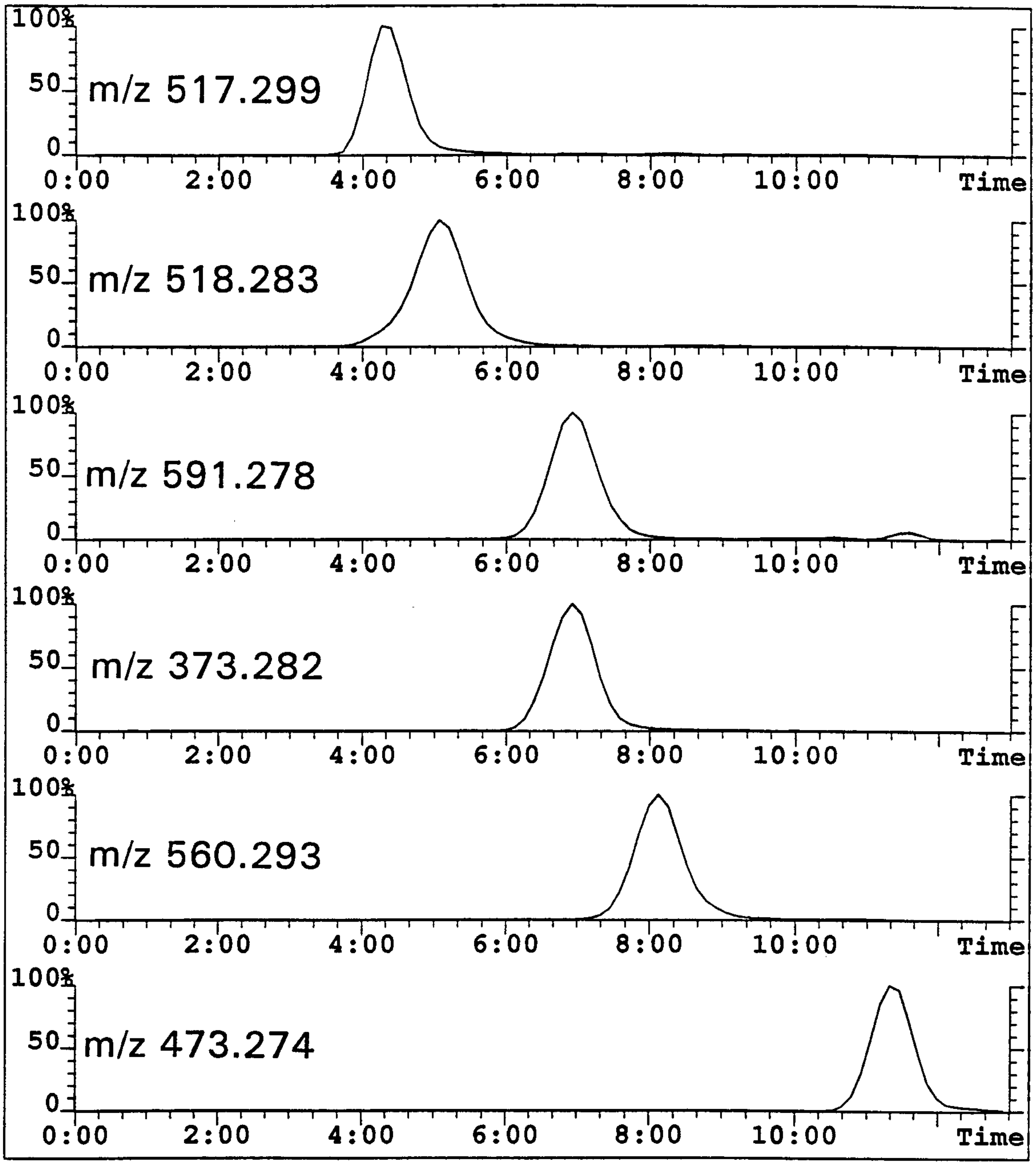

Figure 3: Reconstructed-ion-current chromatograms for the $(M+H)^+$ ions of tryptic fragments of synthetic peptide 18 obtained during LC-ESI-MS analysis with a sector resolution of 3000 (10% valley definition) and a sampling cone voltage of 75 volts.

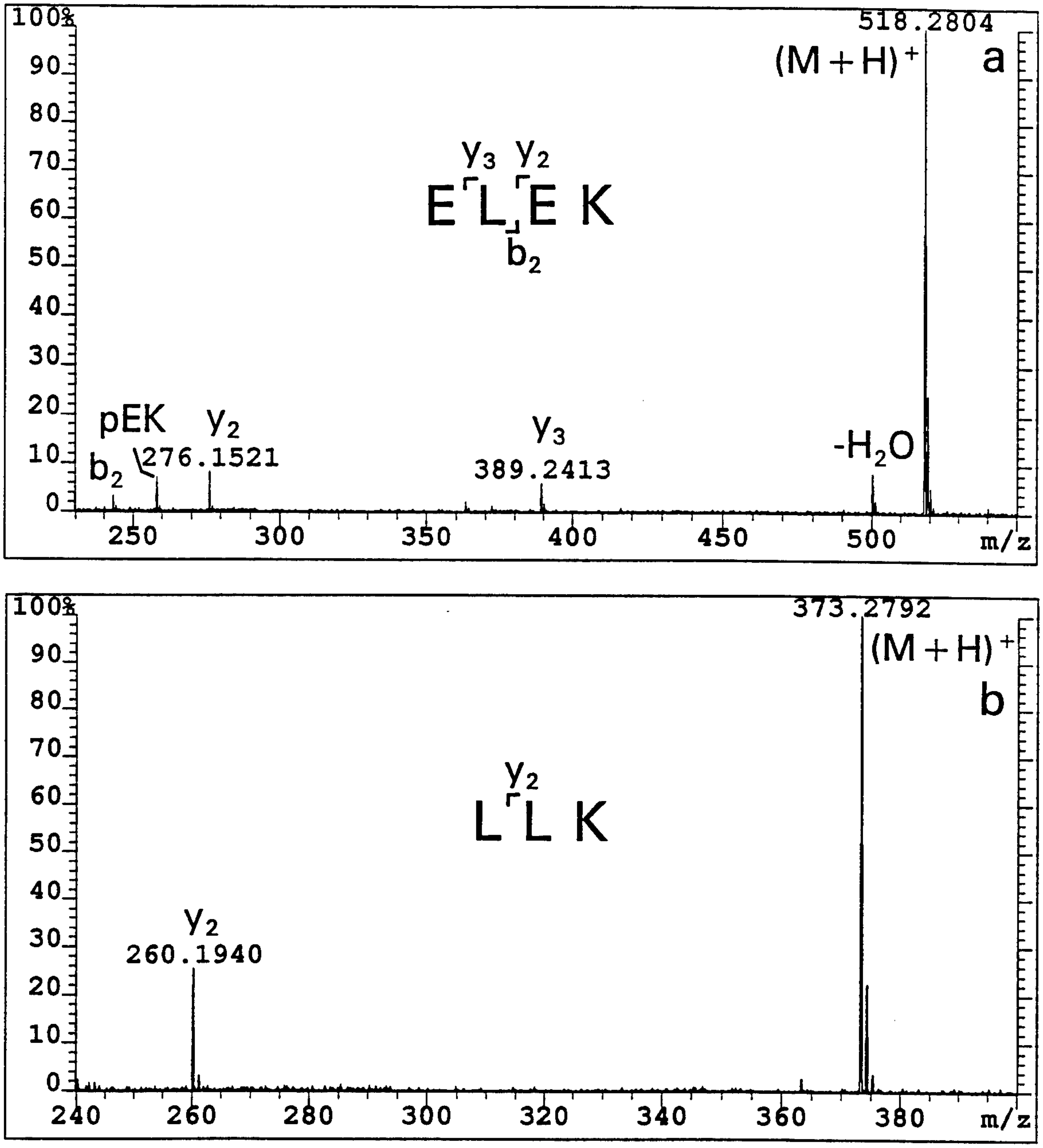

Figure 4: Typical CAD/MS data obtained for tryptic fragments a) ELEK and b) LLK obtained during LC-ESI-MS analysis with a sector resolution of 3000 (10% valley definition) and a sampling cone voltage of 150 volts.

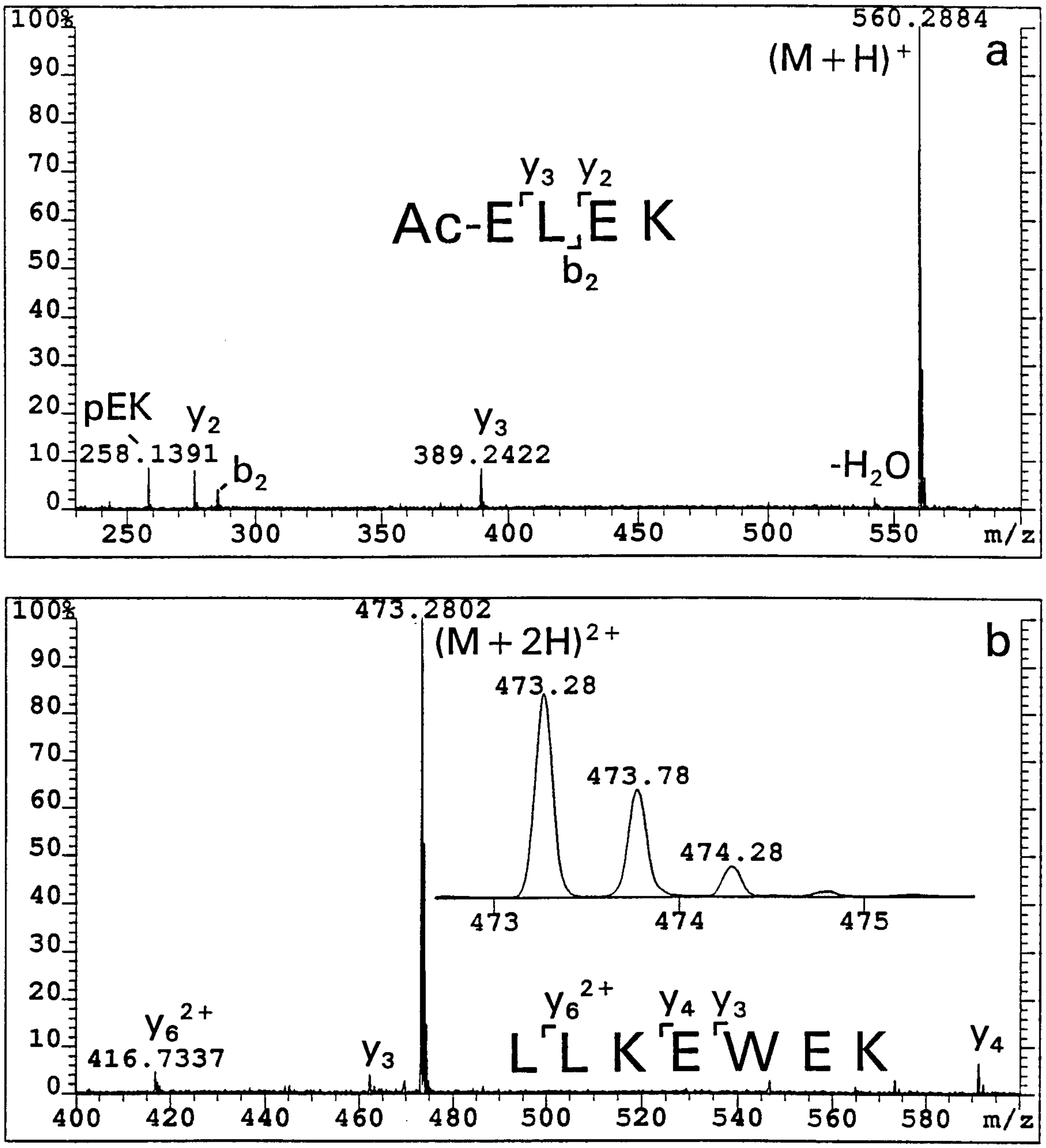

Figure 5: Typical CAD/MS data obtained for tryptic fragments a) Ac-ELEK and b) LLKEWEK obtained during LC-ESI-MS analysis with a sector resolution of 3000 (10% valley definition) and sampling cone voltages of 150 and 115 volts, respectively.

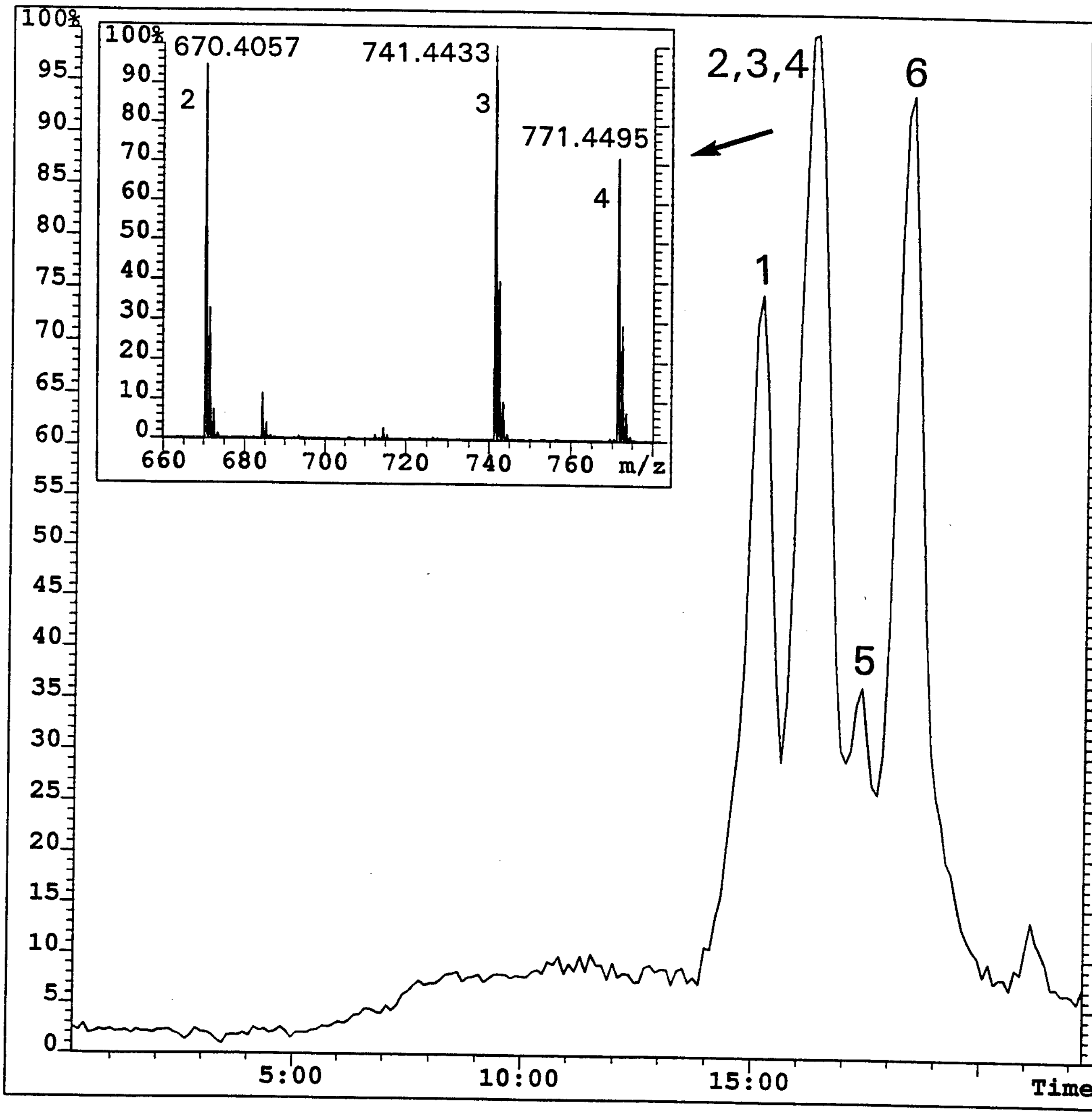

Figure 6: LC-ESI-MS total-ion-current (910 to 575 Da) chromatogram of the peptides in an incomplete synthetic reaction mixture obtained with a resolution of 3500 (10% valley definition) and a sampling cone voltage of 75 volts. (Inset: $(M+H)^+$ isotopic clusters for peptides 2, 3 and 4).

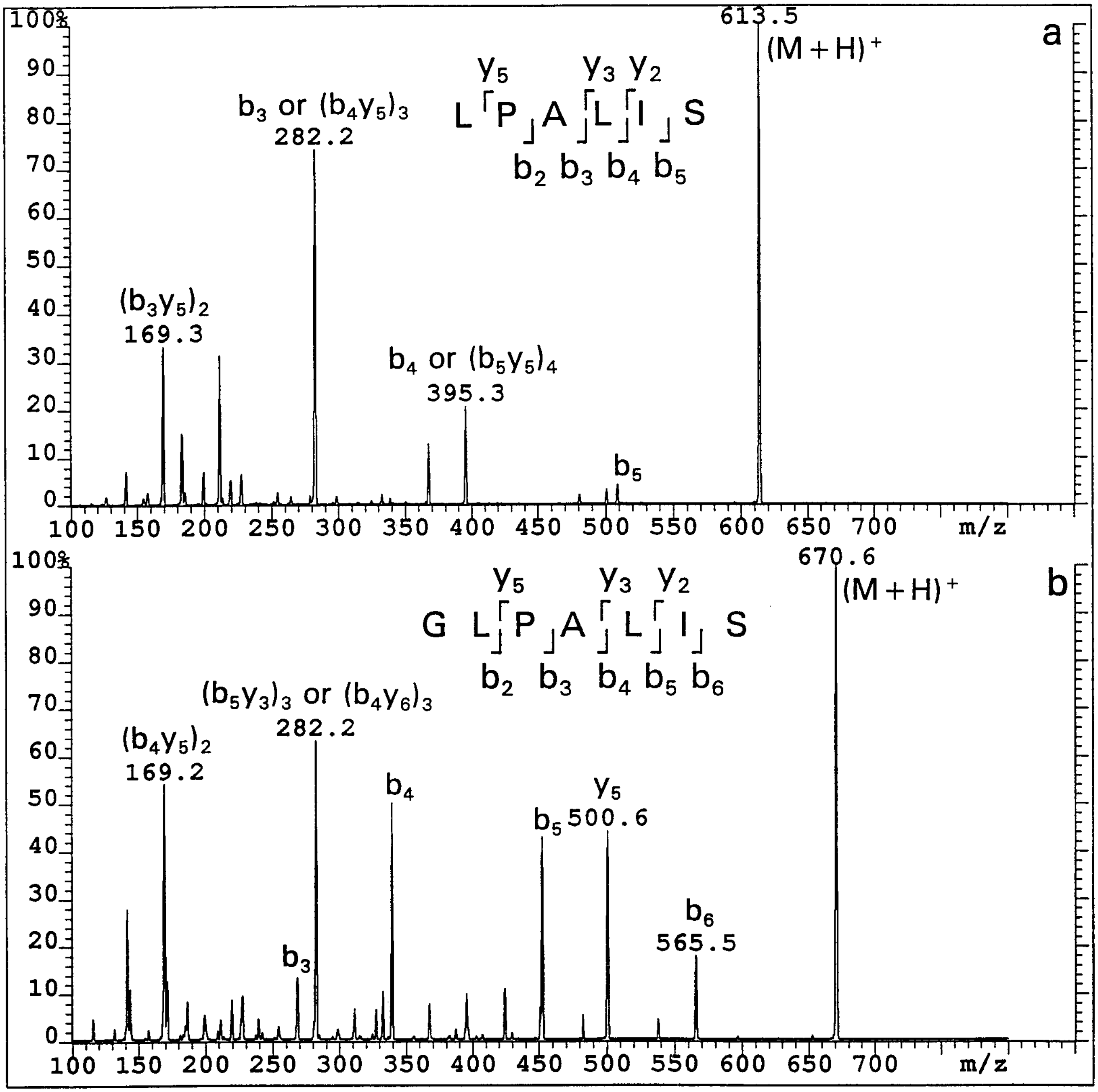

Figure 7: Product ion mass spectra for peptides a) LPALIS and b) GLPALIS obtained during loop injection of the incomplete synthetic reaction mixture. (Quadrupole CAD cell: 100 volts, 1.1×10^{-4} Torr argon).

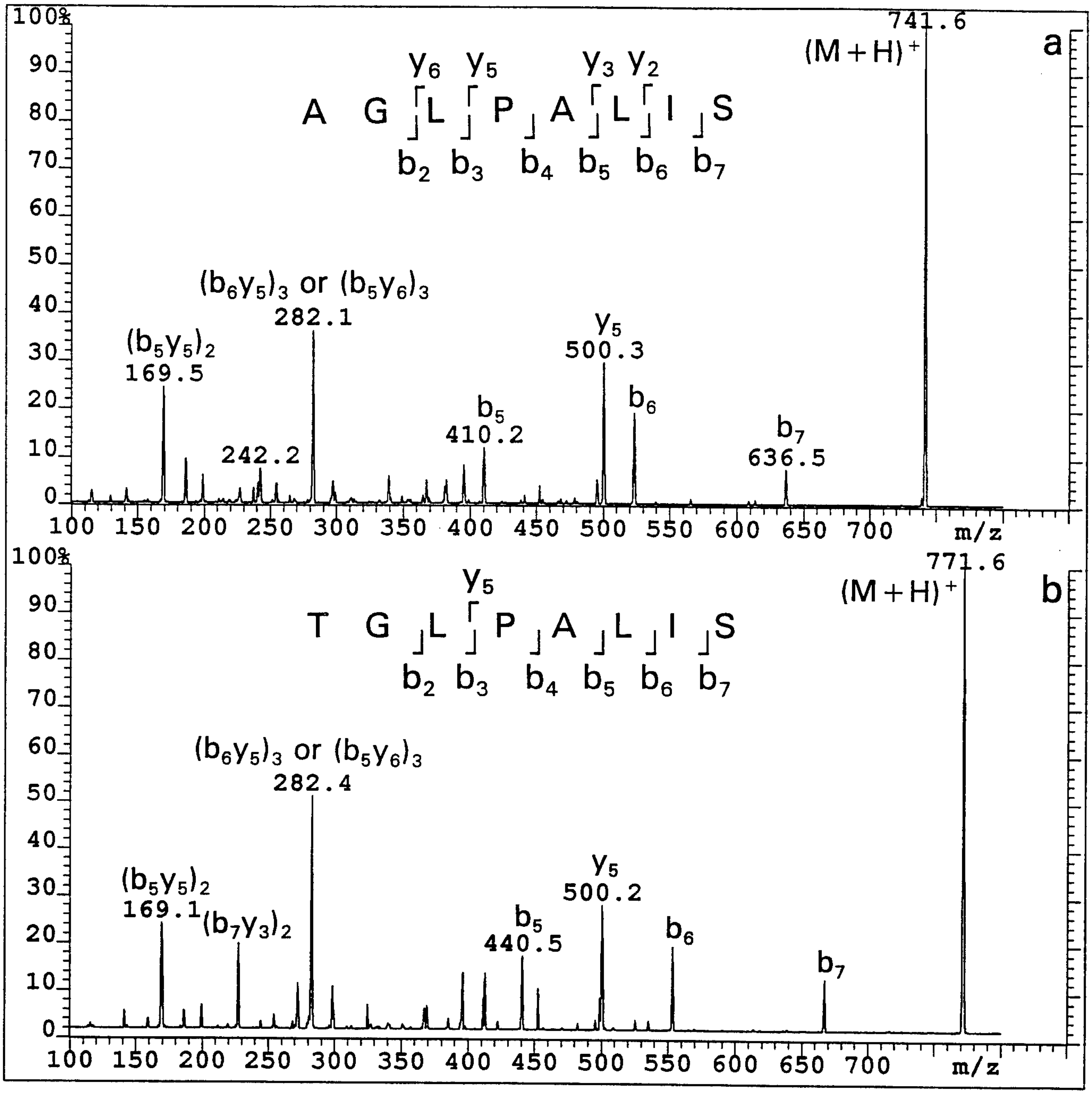

Figure 8: Product ion mass spectra for peptides a) AGLPALIS and b) TGLPALIS obtained during loop injection of the incomplete synthetic reaction mixture. (Quadrupole CAD cell: 100 volts, 1.1×10^{-4} Torr argon).

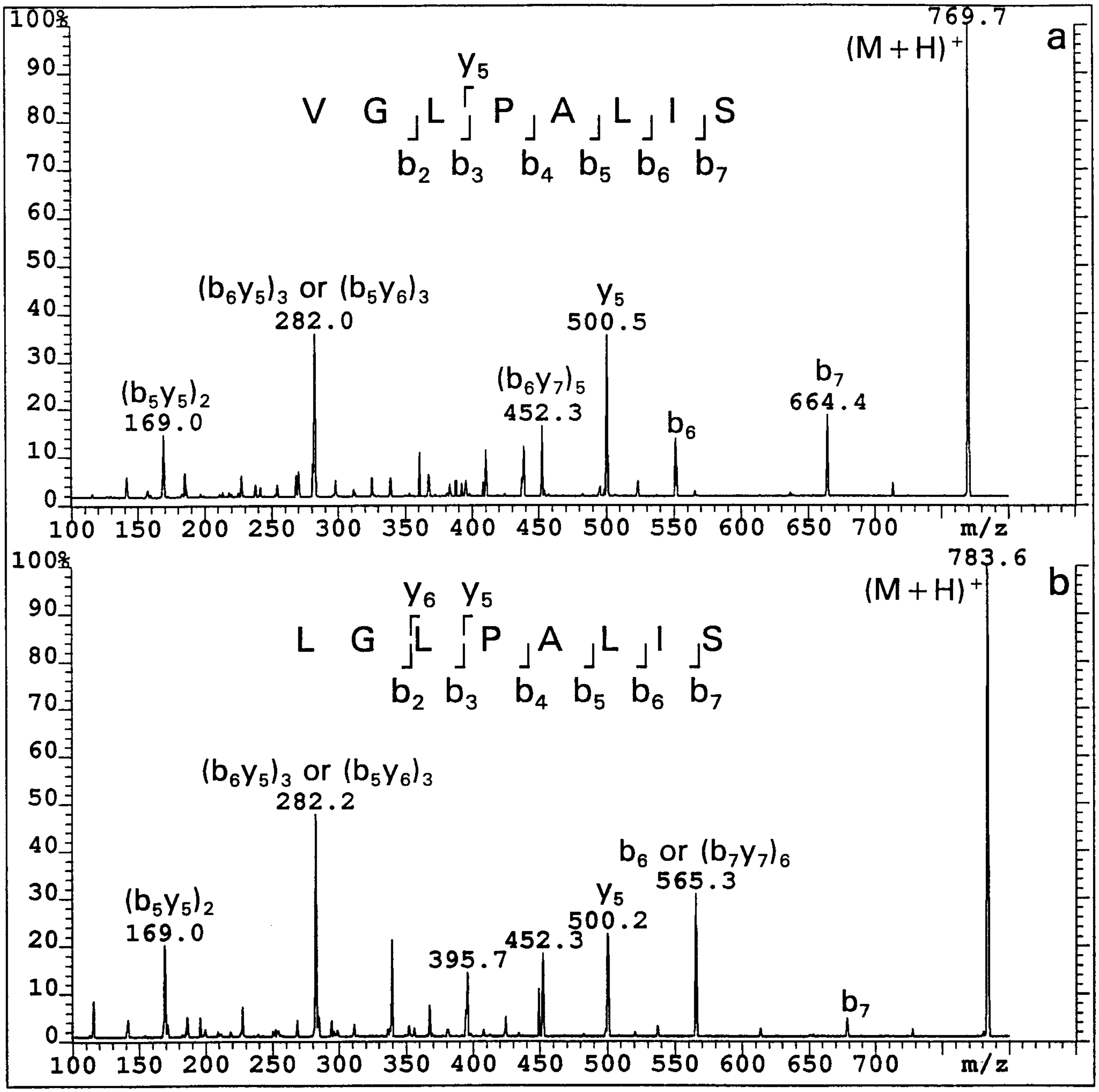

Figure 9: Product ion mass spectra for peptides a) VGLPALIS and b) LGLPALIS obtained during loop injection of the incomplete synthetic reaction mixture. (Quadrupole CAD cell: 100 volts, 1.1×10^{-4} Torr argon).

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence Research Establishment Suffield
Box 4000, Medicine Hat, AB T1A 8K6

2. SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable)

Unclassified

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title.)

Liquid Chromatographic Mass Spectrometric and Tandem Mass Spectrometric Identification of Peptides Using Electrospray Ionization

4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.)

Hancock, J.R., D'Agostino, P.A. and Provost, L.R. (DRES)
Semchuk, P.D. and Hodges, R.S. (University of Alberta)

5. DATE OF PUBLICATION (month and year of publication of document)

November 1996

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

33

6b. NO. OF REFS (total cited in document)

23

6. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Suffield Report

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

SR 638

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availabilty (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

High resolution liquid chromatography electrospray mass spectrometry (LC-ESI-MS) was investigated for the identification of known and unknown peptides in a research effort designed to evaluate the applicability of this and complementary MS techniques for the characterization and identification of peptides. The monoisotopic molecular weights of five related peptides with molecular weights between 2000 and 2500 Da, typical of bioactive peptides, were acquired with a magnetic sector resolution of 3000 (10% valley definition). Tryptic maps were generated for each peptide during high resolution LC-ESI-MS analysis and collisionally activated dissociation (CAD) in the ESI interface resulted in the production of characteristic product ions that enabled amino acid sequencing of the tryptic fragments. Under narrow and wide mass range magnetic sector scanning conditions monoisotopic molecular weight errors were typically in the 10-20 and 30-40 ppm range, respectively. Unknown identification, a critical requirement for MS methods developed in support of chemical/biological sample analysis, was demonstrated during analysis of an incomplete synthetic peptide reaction mixture. The reaction mixture for the peptide, LTTAVKKVLTTGLPALIS, was selected since the 18 amino acid target peptide was not synthesized. In its place were a number of lower molecular weight unknown peptides. The monoisotopic molecular weights of these unknowns was determined to within 10-20 ppm with a magnetic sector resolution of 3500 (10% valley definition). However, co-elution of the unknown peptides during LC-ESI-MS analysis precluded the use of CAD-MS for amino acid sequencing. Amino acid sequence information for six unknown peptides was generated by ESI-MS/MS, making use of the quadrupole collisional cell associated with the hybrid tandem mass spectrometer. The product spectra acquired for the $(M+H)^+$ precursor ions for these peptides were dominated by b_n-series ions from which the amino acid sequence could be assigned. Supplementary information on the amino acid sequence was provided by the presence of y_n-series and internal fragment ions.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Peptide

Mass Spectrometry

Tandem Mass Spectrometry

High Resolution

Liquid Chromatography

Electrospray